

Elements of kinetics

Statistical mechanics does not tell us anything about how fast things happen. For example, from equilibrium considerations, wood, in the presence of oxygen, should turn into carbon dioxide and water at room T. It however doesn't unless we burn it (i.e., increase the temperature) given typical observation times.

Consider an isomerization process



We know from experience, that the amounts of A and B are well described by phenomenological equations:

$$\frac{dN_A}{dt} = -k_f N_A + k_B N_B$$
$$\frac{dN_B}{dt} = k_f N_A - k_B N_B$$

We note that these equations are not exact: they are accurate enough only if we are looking at time scales that are slow enough: If we study molecules at fast (say picosecond or femtosecond time scales) then these equations will not be correct.

In equilibrium, the lhs of these eqs. should be zero so that

$$k_f N_A = k_B N_B$$

On the other hand,

$$N_B/N_A = q_B/q_A = K,$$

where q_A and q_B are the partition functions corresponding to A and B and K is the equilibrium constant. Therefore we expect

$$k_f q_A = k_B q_B$$

I.e., the ratio of the two rate constants, k_f and k_B should be equal to the ratio of the partition functions q_B/q_A equal to the equilibrium constant of the reaction.

If we divide our kinetic equations by the total number of atoms N we will find:

$$\frac{dp_A}{dt} = -k_f p_A + k_b p_B$$

$$\frac{dp_B}{dt} = k_f p_A - k_b p_B$$

where p_A and p_B are the probabilities for the molecule to be in the forms A and B.

Suppose that at we prepared at $t=0$ a non-equilibrium ensemble of molecules such that a fraction $p_A(0)$ of the molecules is in state A and $p_B(0) = 1 - p_A(0)$ in state B. From probability theory, we know that

$$p_A(t) = p(A, t | A, 0) p_A(0) + p(A, t | B, 0) p_B(0)$$

$$p_B(t) = p(B, t | A, 0) p_A(0) + p(B, t | B, 0) p_B(0)$$

Here $p(B, t | A, 0)$ is the conditional probability that one finds the molecule in state B at time t provided that it was in state A at $t=0$ etc. This follows from the fact that $p(A, t | B, 0) p_B(0)$ is the joint probability of finding the molecule in state B at $t=0$ and in state A at time t etc. We can write the above equation as matrix multiplication:

$$\begin{pmatrix} p_A(t) \\ p_B(t) \end{pmatrix} = \begin{bmatrix} p(A, t | A, 0) & p(A, t | B, 0) \\ p(B, t | A, 0) & p(B, t | B, 0) \end{bmatrix} \begin{pmatrix} p_A(0) \\ p_B(0) \end{pmatrix}$$

On the other hand, if we solve the above kinetic equations directly, we will find:

$$\begin{pmatrix} p_A(t) \\ p_B(t) \end{pmatrix} = \begin{bmatrix} \frac{k_b}{k_b + k_f} + \frac{k_f}{k_b + k_f} e^{-kt} & \frac{k_b}{k_b + k_f} (1 - e^{-kt}) \\ \frac{k_f}{k_b + k_f} (1 - e^{-kt}) & \frac{k_f}{k_b + k_f} + \frac{k_b}{k_b + k_f} e^{-kt} \end{bmatrix} \begin{pmatrix} p_A(0) \\ p_B(0) \end{pmatrix}$$

where $k = k_b + k_f$.

For $t \rightarrow \infty$ we find

$$p_A(\infty) = \frac{k_b}{k_b + k_f} (p_A(0) + p_B(0)) = \frac{k_b}{k_b + k_f}$$

$$p_B(\infty) = \frac{k_f}{k_b + k_f}$$

so that $k_f p_A(\infty) = k_b p_B(\infty)$ as we expect in equilibrium.

You can also check that the sum of the probabilities p_A and p_B is equal to 1 for any t .

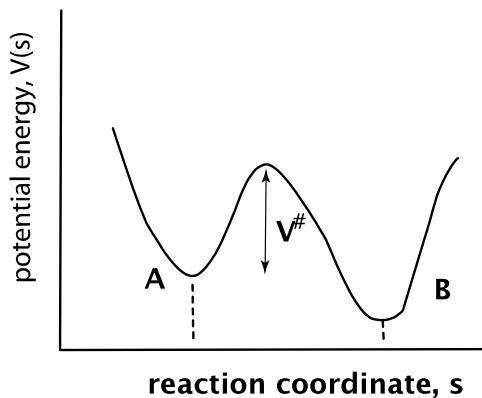
Another result that will prove useful is that

$$k_f = \left. \frac{d}{dt} p(B, t | A, 0) \right|_{t \rightarrow 0}$$

Loosely speaking, k_f is the probability of going from A to B per unit time. Similarly,

$$k_b = \left. \frac{d}{dt} p(A, t | B, 0) \right|_{t \rightarrow 0}$$

Now we'd like to relate the above phenomenological description to a more microscopic picture of the dynamics. Suppose, for example, that the dynamics of our system can be described as one-dimensional motion along a single reaction coordinate s :



Our goal is to calculate the rate constants k_f and k_b from the properties of the potential energy profile $V(s)$. We will assume that the motion of our system obeys the classical-mechanical equations of motion:

$$m\ddot{s} = -\frac{\partial V}{\partial s}$$

We could in principle solve the problem by initially creating an ensemble of molecules with different initial conditions, $(s_i(0), p_i(0) = m\dot{s}_i(0))$. We can assume that our ensemble obeys the Boltzmann statistics. That is, the probability density for the variables $s_i(0), p_i(0)$ is

$$w(s_i(0), p_i(0)) \propto \exp(-\beta(V(s_i(0)) + p_i^2(0)/2m))$$

We now let our ensemble of trajectories evolve independently, each starting with its own initial condition. If at a time t we examine our ensemble and measure the coordinate and the momentum, $s_i(t), p_i(t)$, we will discover that our ensemble still obeys the Boltzmann statistics (this is because the energy and the phase-space volume are conserved along the trajectories).

The correct Boltzmann distribution is the equilibrium distribution for our system. This distribution will not change with time. We can however imagine that we have initially excluded any initial conditions with $s > s^\ddagger$. In other words, we have prepared an ensemble of molecules A only. Now we can watch our molecules undergo the reaction $A = B$. If we follow the time evolution of each trajectory we then discover that some trajectories will cross over to the right of the dividing line $s=s^\ddagger$ thus transforming into B. If we count the number $N_B(t) = Np_B(t)$ of the trajectories that have $s(t) > s^\ddagger$, we then can substitute this result into our kinetic equations and from them calculate k_f and k_b .

In practice, the above plan is typically doomed. The problem is that a trajectory crossing over the barrier at $s = s^\ddagger$ is a very unlikely event. Indeed, for that the molecule should have an extra energy of at least V^\ddagger . The number of such molecules is thus proportional to the factor $\exp(-V^\ddagger/k_B T)$. Since typically $V^\ddagger > k_B T$, the fraction of such successful trajectories is very small. If, for example, $\exp(-V^\ddagger/k_B T) \sim 10^{-10}$ and we generate say 1 million trajectories on the computer, then it is very likely that none of them will succeed surmounting the barrier V^\ddagger . Thus we could conclude from our simulation that there is no reaction $A = B$ taking place, which, of course, is wrong.

Intuitively, the rate of crossing from A to B should be proportional to the number of successful trajectories, i.e., $k_f \sim \exp(-V^\ddagger/k_B T)$.

We now show how one can overcome the above difficulty. We use the formula:

$$k_f = \left. \frac{d}{dt} p(B, t | A, 0) \right|_{t \rightarrow 0} = p_A^{-1} \left. \frac{d}{dt} J(B, t | A, 0) \right|_{t \rightarrow 0}$$

where $J(B, t | A, 0)$ is the *joint* probability for being in state A at time 0 and at B at time t.

Introduce the functions

$$\theta_A(s) = \theta(s^\ddagger - s) = \begin{cases} 1, & \text{if } s < s^\ddagger \\ 0, & \text{if } s \geq s^\ddagger \end{cases}$$

$$\theta_B(s) = 1 - \theta_A(s) = \theta(-s^\ddagger + s) = \begin{cases} 1, & \text{if } s \geq s^\ddagger \\ 0, & \text{if } s < s^\ddagger \end{cases}$$

The function $\theta_A(s)$ is equal to 1 if we are on the A side of the potential energy profile and zero otherwise. The purpose of introducing such a function is to map the continuous coordinate s onto a discrete, two-state system that has only two states, A and B.

Consider the quantity

$$\langle \theta_B(s(t))\theta_A(s(0)) \rangle = q^{-1} \int \frac{d\Gamma}{2\pi\hbar} \exp(-\beta H(\Gamma)) \theta_A(s(0))\theta_B(s(t))$$

where Γ denotes a point in the phase space and

$$H(\Gamma) = \frac{p^2(0)}{2m} + V(s(0))$$

Quantities of the above kind are called correlation functions.

Before we go on with our derivation, we need to discuss how integrals can be evaluated by Monte Carlo method. Suppose we want to calculate the average $\langle f(x) \rangle$ given that x has a normalized probability distribution $W(x)$. We have

$$\langle f(x) \rangle = \frac{\int_{-\infty}^{\infty} f(x)W(x)dx}{\int_{-\infty}^{\infty} W(x)dx} = \int_{-\infty}^{\infty} f(x)W(x)dx$$

We could do it directly by evaluating the above integral numerically. However in many cases (especially when such an integral needs to be evaluated in more than one dimension) Monte Carlo importance sampling methods are more advantageous. Suppose that our computer has a random number generator that can generate random numbers x_i , $i = 1, 2, \dots, N$, with the probability distribution $W(x)$. Then in the limit $N \rightarrow \infty$ the above integral can be calculated as:

$$\langle f(x) \rangle = \frac{1}{N} \sum_{i=1}^N f(x_i).$$

To see this we note that the number of x_i 's that will be generated in the range $x < x_i < x + \Delta x$ is equal to $NW(x) \Delta x$ so that the above sum can be written as

$$\frac{1}{N} \sum_{i=1}^N f(x_i) \approx \frac{1}{N} \sum_x NW(x) f(x) \Delta x \approx \int_{-\infty}^{\infty} dx f(x) W(x)$$

We are going to calculate the above correlation function using Monte Carlo as follows: We consider an equilibrium Boltzmann ensemble of trajectories, with positions and momenta distributed according to the Boltzmann distribution. As argued above, once Boltzmann, the ensemble stays Boltzmann. If, at time $t=0$, we look at the positions and the momenta of each member of the ensemble, we will find that they obey the Boltzmann distribution. So we can generate our ensemble by launching trajectories with the initial position and momentum, $(s(0), p(0))$, taken from the Boltzmann distribution:

$$W(s(0), p(0)) dp(0) ds(0) = W(\Gamma) d\Gamma = q^{-1} \frac{d\Gamma}{2\pi\hbar} \exp(-\beta H(\Gamma))$$

We then follow each trajectory over the period of time equal to t . If $s(0)$ is to the left of s^\ddagger and $s(t)$ is to the right of s^\ddagger then the contribution from such a trajectory will be equal to 1. Otherwise the trajectory will contribute zero due to the presence of the functions θ_A and θ_B in the integrand. In other words, the product $\theta_A(s(0))\theta_B(s(t))$ will select the trajectories that start, at $t=0$, on the reactant (A) side of the barrier and end up, at time t , on the product side. Therefore $\langle \theta_A(s(0))\theta_B(s(t)) \rangle$ is the fraction of such trajectories in the Boltzmann ensemble of the trajectories. In other words, it is the joint probability $P(B, t | A, 0)$ that we are in state A at $t=0$ and state B at time t . To get the conditional probability $p_{BA}(t)$, we simply need to divide the above quantity by the equilibrium probability p_A that we are in state A at $t=0$ (recall from probability theory that the joint probability of two events, $p(b, a)$ is the probability of a times the conditional probability $p(b|a)$ for b to happen given that a happened.)

We have

$$p(B, t | A, 0) = \langle \theta_A(s(0))\theta_B(s(t)) \rangle / p_A$$

where $p_A = \langle \theta_A(s(0)) \rangle = q_A/q$.

We found

$$p(B, t | A, 0) = \langle \theta_A(s(0)) \theta_B(s(t)) \rangle / p_A$$

where $p_A = \langle \theta_A(s(0)) \rangle = q_A/q$.

$$k_f(t) = dp_{BA}(t) / dt = \langle \theta_A(s(0)) \dot{\theta}_B(s(t)) \rangle / p_A = - \langle \dot{\theta}_A(s(0)) \theta_B(s(t)) \rangle / p_A = \langle \dot{\theta}_B(s(0)) \theta_B(s(t)) \rangle / p_A$$

We have used the fact that $\theta_A = 1 - \theta_B$ and that

$$\langle A(0) \dot{B}(t) \rangle = - \langle \dot{A}(0) B(t) \rangle$$

The latter property results from time-translation symmetry of classical trajectories, which implies

$$\langle A(0) B(t) \rangle = \langle A(-t) B(0) \rangle$$

Differentiating this with respect to t gives the above identity.

Using the formula

$$\frac{d}{dt} \theta(s(t) - s^\ddagger) = \dot{s}(t) \delta(s(t) - s^\ddagger)$$

we find

$$k_f = p_A^{-1} \langle \dot{s}(0) \delta(s(0) - s^\ddagger) \theta(s(t) - s^\ddagger) \rangle |_{t \rightarrow 0}$$

The reason why this expression is advantageous numerically is the presence of the δ -function in it: This selects the trajectories that originate at the dividing point s^\ddagger and thus have a fair chance to make transition from A to B.

In the limit $t \rightarrow 0$ we can write $s(t) = \dot{s}(0)t + s(0)$

$$k_f = p_A^{-1} \langle \dot{s}(0) \delta(s(0) - s^\ddagger) \theta(\dot{s}(0)t) \rangle |_{t \rightarrow 0}$$

The presence of $\theta(\dot{s}(0)t)$ in our expression selects trajectories with positive initial velocity, $\dot{s}(0) > 0$. In other words, only the trajectories going from left to right (A to B) contribute to k_f . We can rewrite our expression as

$$\begin{aligned} k_f &= p_A^{-1} q^{-1} \iint_{p>0} \frac{dp ds}{2\pi\hbar} \left(\frac{p}{m} \right) \delta(s - s^\ddagger) \exp\left(-\beta \left[\frac{p^2}{2m} + V(s) \right]\right) \\ &= q_A^{-1} \exp(-\beta V(s^\ddagger)) \int_0^\infty \frac{dp}{2\pi\hbar} \left(\frac{p}{m} \right) \exp\left(-\beta \frac{p^2}{2m}\right) = \frac{k_B T}{2\pi\hbar q_A} \exp(-\beta V(s^\ddagger)) \end{aligned}$$

Another way to write this is:

$$k_f q_A = \sqrt{\frac{k_B T}{2\pi m}} q^\ddagger$$

where

$$q^\ddagger = \iint \frac{dp ds}{2\pi\hbar} \delta(s - s^\ddagger) \exp\left(-\beta \left[\frac{p^2}{2m} + V(s) \right]\right)$$

is the constrained partition function at the dividing point. The above expressions for k_f can be interpreted in terms of the average flux of the trajectories crossing the dividing surface (dividing point in our case, but it becomes a hypersurface if we consider more degrees of freedom). The number of such trajectories is proportional to the velocity at the dividing point.

Yet another way to think of our result is

$$k_f = \frac{k_B T}{2\pi\hbar} \frac{q^\ddagger}{q_A}$$

where q^\ddagger is the partition function of the “activated complex”, which is obtained by constraining our system to sit at $s = s^\ddagger$ and excluding integration over the degrees of freedom (p and s) of the unstable mode of the activated complex [in one dimension, there are no other modes so $q^\ddagger = \exp(-\beta V(s^\ddagger))$].

What we have derived is the famous transition state theory (TST) approximation for the reaction rate (We will discuss later why this is an *approximation*).

We can calculate k_b the same way. We will find that

$$k_f q_A = k_b q_B$$

so that the ratio of the two rate constants is, indeed, equal to the equilibrium constant.

It is a bit disconcerting that our classical expression contains Planck's constant. We should however remember that the partition functions $q_{A,B}$ also contain Planck's constant. Typically, it will cancel out in the classical limit.

Suppose, for example, that the motion in the potential well corresponding to molecule A is well described in the harmonic approximation. If the frequency of this motion is ω_A then

$$q_A \approx \frac{k_B T}{\hbar \omega_A} \exp(-\beta V_A),$$

where V_A is the potential energy at the minimum corresponding to A. This gives

$$k_f = \frac{\omega_A}{2\pi} \exp(-\beta(V(s^\ddagger) - V_A))$$

This formula is intuitively appealing: it says that the rate constant is equal to an "attempt frequency" $\omega_A/2\pi$ times the Boltzmann probability to reach the top of the barrier

Recrossings and variational transition state theory

There is a problem with the TST rate: It depends on the location of the point dividing the reactants and products. Experimentally, we know that it shouldn't as our measurements typically don't even know about this point.

To see what went wrong let's examine our original expression for the rate constant:

$$k_f = p_A^{-1} \langle \dot{s}(0) \delta(s(0) - s^\ddagger) \theta(s(t) - s^\ddagger) \rangle_{t \rightarrow 0}$$

This equation says: pick a trajectory at $t=0$ starting at s^\ddagger . The number of such trajectories crossing the dividing surface is proportional to $\dot{s}(0)$. However they should be counted only if they are "reactive", i.e., if they end up on the product side of the barrier. This condition is enforced by the step function $\theta(s(t) - s^\ddagger)$, which excludes any trajectories that are to the left of the dividing point at time t . We have said that the trajectories moving to the right at $t=0$ are reactive. This is however not so because some of them can recross. When we took the $t \rightarrow 0$ limit, we have neglected these recrossings, which seems unphysical. Let's see what happens if we don't impose the $t \rightarrow 0$ limit. For our one-

dimensional potential consider what happens when the dividing point is to the left of the point s_{\max} where $V(s)$ is maximum, $s^\ddagger < s_{\max}$. The trajectories that go from left to right at $t=0$ but do not have enough energy to surmount the rest of the barrier will bounce back. They are not reactive. Thus we modify our expression:

$$k_f = q_A^{-1} \exp(-\beta V(s^\ddagger)) \int_{p_{\min}}^{\infty} \frac{dp}{2\pi\hbar} \left(\frac{p}{m} \right) \exp\left(-\beta \frac{p^2}{2m}\right)$$

where p_{\min} is the minimum momentum a trajectory must have initially so that it can overcome the barrier. This can be found from the equation:

$$p_{\min}^2 / 2m + V(s^\ddagger) = V(s_{\max})$$

Performing the integral then gives us:

$$k_f = \frac{k_B T}{2\pi\hbar q_A} \exp(-\beta V(s_{\max})),$$

a result independent of s^\ddagger .

The correct correlation function recipe to calculate k_f is as follows:

$$k_f = p_A^{-1} \langle \dot{s}(0) \delta(s(0) - s^\ddagger) \theta(s(t) - s^\ddagger) \rangle_{|_{t_m < t < t_r}} \quad (*)$$

The correlation function is to be evaluated for times that are longer than the molecular time scale t_m , which is roughly the time scale for the trajectories to recross, but shorter than the time scale t_r associated with the reaction itself (i.e. $\sim k_f^{-1}$). In this latter sense this is still a “zero-time limit”: the molecular time scale is not resolved by our phenomenological rate equations that depend only on k_f and k_b .

From the above arguments, transition state theory always overestimates the rate constant because it assumes that every trajectory launched towards the products is reactive. One thus can come up with a variational transition-state theory, in which one would estimate an upper bound on k_f

$$k_f = \frac{k_B T}{2\pi\hbar q_A} \exp(-\beta V(s^\ddagger))$$

with respect to the location of the dividing plane s^\ddagger . One can see that in our particular simple one-dimensional case this will give the exact answer, with the optimum plane located at s_{\max} .

Note that equation (*) is more general than the model adopted above: It does not in fact assume Newtonian dynamics along the coordinate s . Instead, our system can contain any number of degrees of freedom, while s can represent a one-dimensional order parameter (or experimental signal) used to monitor the reaction progress. For a commonly used model, where the dynamics along s is stochastic and is described by a Langevin equation (which, in addition to the conservative force $-dV/ds$, involves a velocity dependent force and a random force both resulting from the interactions of the system of interest with its environment), the rate can also be computed using Eq. (*).